PROCEEDINGS OF TDWG

Biodiversity Information Science and Standards 1: e20236
doi: 10.3897/tdwgproceedings.1.20236

OPEN ACCESS

Conference Abstract

What's Missing From All the Portals?

Sharon Grant‡, Kate Webbink‡, Janeen Jones‡, Pete Herbst‡, Robert Zschernitz‡, Rusty Russell‡

‡ Field Museum, Chicago, United States of America

Corresponding author: Sharon Grant (sgrant@fieldmuseum.org), Kate Webbink (kwebbink@fieldmuseum.org)

Received: 12 Aug 2017 | Published: 14 Aug 2017

Citation: Grant S, Webbink K, Jones J, Herbst P, Zschernitz R, Russell R (2017) What's Missing From All the Portals? Proceedings of TDWG 1: e20236. https://doi.org/10.3897/tdwgproceedings.1.20236

Abstract

At time of writing there are over 784 million occurrence records in the Global Biodiversity Information Facility (GBIF) portal (gbif.org), 106 million on the iDigBio site (idigbio.org); 68 million in the Atlas of Living Australia (ala.org.au) and 20 million in VertNet (vertnet.org). The list of biodiversity aggregators and portals that boast occurrence counts in the millions continues to increase. Combined with sites who gather data their data from outside of the GBIF domain such as The Paleobiology Database, there is compelling evidence that global digitization is starting to illuminate the black hole of biodiversity data held in collections across the world. The visibility and demands on our collective natural history heritage have never been as high, and they are increasingly in the spotlight with both internal and external audiences. Funding sources have moved away from massive "digitization for the sake of digitization" projects and demand much more focused proposals. To compete in this arena, collections staff and researchers must collaborate and mine collections for their strengths and use those to justify efforts. To do this, however, they must have access to information about the non-digitized occurrence level records in the world's holdings.

We discuss the potential use of current TDWG standards to allow the capture of existing institutional data about undigitized collections and also those whose records have been marked as environmentally, culturally, or politically sensitive and so must remain digitally dark, so that portals like GBIF can use them in a comparable way as existing occurrence records. Can Darwin Core (with its extensions) together with the Natural Collections Description (draft standard) be used to describe accessions, inventory-level information, and backlog estimates in an efficient and effective way and provide even greater visibility of

© Grant S et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

those undigitized occurrences? In addition, can these data also serve as a means to further refine existing digitized records?

Keywords

occurrence, Darwin Core, backlog, dark data, digitization, Natural Collections Description

Presenting author

Sharon Grant